The Anatomy of an Optical Biopsy Semantic Retrieval System

Ruben Tous, Jaime Delgado, Thomas Zinkl, Pere Toran, and Gabriela Alcalde
Universitat Politecnica de Catalunya, Spain

Martin Goetz
Mainz University Hospital, Germany

Olga Ferrer-Roca
University of La Laguna, Spain

A case-based computer-aided diagnosis system assists physicians and other medical personnel in the interpretation of optical biopsies obtained through confocal laser endomicroscopy.

Medical environments today generate digital images to assist physicians and other health professionals with decision-making tasks. The increasing number of images and usage scenarios, however, necessitates the provision of the proper tools to help medical personnel manage and interpret the images. Computer technology is easing the medical image interpretation process by incorporating computer-aided detection (CADe) and computer-aided diagnosis (CADx) technologies.

One CADx approach includes tools for searching and retrieving information (audiovisual content, text, and metadata) about precedent diagnoses. In the past few years, researchers have extended the functionality of such systems, traditionally based on text-retrieval techniques, by adopting content-based image retrieval (CBIR) techniques. (See the "Related Work in Image Retrieval Systems" for more details.) CBIR lets users search for digital images by relying on those images' low-level visual features, overcoming some of the limitations inherent to text-based medical retrieval systems, such as the difficulty of manually maintaining metadata annotations.

However, traditional low-level features can quantify only basic visual aspects of an image, such as the color histogram. In contrast, human perception of images and image search processes involves high-level abstract formulations, often involving contextual knowledge. The difference between a low-level computational representation and high-level human perception is known as the semantic gap.

This articles proposes a case-based CADx system that assists physicians and other medical personnel in the interpretation of optical biopsies obtained through *confocal laser endomicroscopy* (CLE), which is a novel technique for intravital microscopy during ongoing gastrointestinal endoscopy. However, most gastroenterologists are not trained to interpret mucosal pathology, and histopathologists are generally unavailable in the endoscopy suite. An *optical biopsy* is an optic diagnosis method that can analyze the tissue surface and depth (using a laser or other method) without needing to extract it from the body. The system lets users navigate and search an image database containing optical biopsies of the human colon recorded with CLE. They can then retrieve information about precedent diagnostics by providing an example CLE image for CBIR using keywords or by filtering different fields for structured retrieval. The proposed system's CBIR approach involves an algorithm for automatic feature extraction in CLE images, showing promising results on inferring semantic metadata from low-level features. To effectively ensure the interoperability with potential third-party applications, the system provides an interface compliant with the MPEG Query Format (ISO/IEC 15938-12:2008)[1] and JPEG Search (ISO/IEC 24800)[2] standards. The system's evaluation results are based on test sets of CLE images provided by the Mainz University Hospital in Germany.

Optical Biopsy with CLE

One particular use case of medical CBIR is related to the introduction of new technologies for the microscopic examination of human tissue. In clinical medicine, histopathology is currently the usual procedure for examining the tissue from the human body to study the manifestations of disease. A histopathological examination starts by removing the tissue

1070-986X/12/$31.00 © 2012 IEEE

Published by the IEEE Computer Society

Related Work in Image Retrieval Systems

Researchers in previous work presented an architecture of a complete image management and retrieval system that supports medical tasks such as diagnosis and telemedicine.[1] Their proposed system supports content-based retrieval, but it operates over traditional histological images instead of optical biopsies. Douglas De Macedo and his colleagues described a complete system architecture that stores and retrieves Digital Imaging and Communications in Medicine (DICOM) medical images, focusing on high-performance scalability and information distribution.[2] They defined their own data model, whereas we rely on the ISO/IEC 24800 (JPSearch) standard.[3] Jayashree Kalpathy-Cramer and his colleagues designed and implemented a Web-based retrieval system involved in the medical retrieval task of ImageCLEF 2008.[4] That system was built using a Ruby on Rails framework and Ferret, a Ruby port of Lucene. Our system is also based on Lucene, but it is implemented using Java and specializes in optical biopsies.

Mathias Lux and Savvas Chatzichristofis also proposed a well-known approach based on the use of Lucene.[5] Compared to this work, our approach further extends the Lucene framework with our content-based image retrieval (CBIR) modules while preserving its capability to orchestrate Boolean condition trees. Lux and Chatzichristofis adopted the Lucene interface to process CBIR queries but completely replace the Lucene query evaluation mechanism.

As far as we know, our system differs from any previous approach in that it uses a standard query interface (ISO/IEC 15938-12:2008) capable of expressing complex user search conditions that combine CBIR, metadata, and keywords. Regarding the CBIR part, Barbara André and her colleagues[6] took a different approach and used the bag-of-words (BoW) method.[7] However, their approach is based on another confocal endomicroscopy system, Cellvizio, developed by Mauna Kea Technologies. This approach lets physicians record video during an examination, but it has a smaller field of view and lower resolution than the system we use (Pentax), which lets us see a whole group of crypts in one image.

Henning Mueller and his colleagues provided a general survey of content-based image retrieval systems in medical applications, explaining the propositions for the use of image retrieval in medical practice and the various approaches.[8] One of their conclusions was that, although many propositions for systems are made from the medical domain, research prototypes are developed in computer science departments using medical datasets, and few systems seem to be used in clinical practice. They explained that the goal is not, in general, to replace current text-based retrieval methods but to complement them with visual search tools.

References

1. J.C. Caicedo et al., "Design of a Medical Image Database with Content-Based Retrieval Capabilities," *Proc. 2nd Pacific Rim Conf. Advances in Image and Video Technology,* Springer-Verlag, 2007, pp. 919–931.
2. D.D.J. De Macedo et al., "An Architecture for DICOM Medical Images Storage and Retrieval Adopting Distributed File Systems," *Int'l J. High Performance Systems Architecture,* vol. 2, no. 2, 2009, pp. 99–106.
3. *ISO/IEC 24800-2:2011 Information technology – JPSearch – Part 2: Registration, Identification, and Management of Schema and Ontology,* 2011.
4. J. Kalpathy-Cramer et al., "Multimodal Medical Image Retrieval OHSU at ImageCLEF 2008," *Proc. 9th Cross-language Evaluation Forum Conf. Evaluating Systems for Multilingual and Multimodal Information Access* (CLEF 08), Springer-Verlag, 2008, pp. 744–751.
5. M. Lux and S.A. Chatzichristofis, "Lire: Lucene Image Retrieval: An Extensible Java CBIR Library," *Proc. 16th ACM Int'l Conf. Multimedia* (MM 08), ACM Press, 2008, pp. 1085–1088.
6. B. André et al., "Endomicroscopic Image Retrieval and Classification Using Invariant Visual Features," *Proc. 6th IEEE Int'l Symp. Biomedical Imaging: From Nano to Macro* (ISBI 09), IEEE Press, 2009, pp. 346–349.
7. J. Yang et al., "Evaluating Bag-of-Visual-Words Representations in Scene Classification," *Proc. Int'l Workshop Multimedia Information Retrieval* (MIR 07), ACM Press, 2007, pp. 197–206.
8. H. Mueller et al., "A Review of Content-Based Image Retrieval Systems in Medical Applications: Clinical Benefits and Future Directions," *Int'l J. Medical Informatics,* vol. 73, no. 1, 2004, pp. 1–23.

from the body (through surgery, biopsy, or autopsy). Then, after preparing the tissue (for example, using chemical fixation), a pathologist examines it under a microscope. Some of the inconveniences and costs associated with this diagnosis method can be overcome using the nonintrusive optical biopsy diagnosis method.

CLE, which this article focuses on, is a new endoscopic modality developed to obtain high-resolution images of the mucosal layer of the gastrointestinal tract in approximately 1,000-fold magnification in vivo after the application of a fluorescent agent such as fluorescein or acriflavine. Typically, the confocal microscope

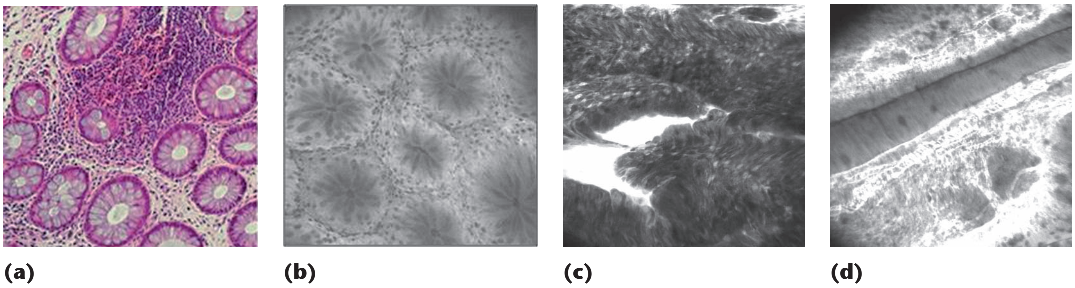

Figure 1. *Example histological and confocal laser endomicroscopy (CLE) images of the colon mucosa. (a) Histology (inflammatory), (b) CLE (healthy) (c) CLE (adenoma), and (d) CLE (carcinoma).*

is integrated into a conventional upper endoscope (we use the Pentax R EC-3870CILK device), and images are captured in vivo during an endoscopic session and digitally stored as grayscale images. Fifty to 500 images from different sites are usually captured per examination. CLE is based on tissue illumination with a low-power laser, with subsequent detection of the fluorescence light rejected from the tissue through a pinhole. The term *confocal* refers to the alignment of both illumination and collection systems in the same focal plane. This notably increases the resolution of confocal endomicroscopy, thus providing an ''optical biopsy'' below the tissue surface.

CLE adoption depends on a thorough knowledge of mucosal architecture and pathology.[3] On the one hand, CLE images and traditional histological images present some visual differences (see Figure 1). Although not substantial, these differences are enough to disallow the usage of the established medical Criterion Standards, which rely on histological images.[4] Before new Criterion Standards based on CLE are established, pathologists need reference material to gain diagnosis confidence through analogy. On the other hand, CLE images are taken and analyzed by endoscopists, who are not generally trained in microscopic morphology, the domain of surgical pathology. This complicates the endoscopists' task of discriminating and annotating the images showing potential pathologies to obtain an immediate preliminary diagnosis or just to simplify the task of the pathologist. For example, for Barrett's esophagus (a premalignancy of the upper gastrointestinal tract), the medical community has deemed it necessary for endoscopists to perform 100 examinations after initial supervised training to be able to reliably diagnose intraepithelial neoplasia during the endoscopy.[5]

To ease such problems, endoscopists could gain classification/diagnosis confidence by locating precedent diagnostics for CLE images with similar features. Our goal is not to avoid the need for a biopsy by a pathologist, but to address real-world situations in which endoscopists working with CLE must select and classify relevant images, generally without the proper training on this kind of microscopic images. The main goal of the current system is facilitating and shortening endoscopists' training. In addition, there is a strong scientific need to explore databases of images for retrospective analyses. However, automatically computing the similarity of two CLE images is a nontrivial task that falls within the CBIR field.

System Information Model

The proposed architecture lets users navigate and search an optical biopsy image database. The multimedia information retrieval (MIR) process usually starts with end users expressing their information needs through a user-friendly query interface. The user's information needs come from the conceptual level and combine criteria about the information represented by the content's data, such as the content's meaning (for example, images showing colonic inflammation) with other criteria about the features of the content's data (for instance, images taken on 3 October). In the end, the only way of fulfilling the user information needs is translating these criteria into machine-readable conditions, as precise as possible, over the content's data, using the data's media binary representation or some metadata annotations. Metadata annotations provide information about the content at different levels, from low-level features and management information to semantic-level descriptions.

Thus, the problem is twofold. On the one hand, we have the challenge of enriching the media data with metadata useful for solving the queries. On the other hand, we must formalize the user information needs, as much as possible, in the form of a query expressed in terms of the available metadata model. Even if we succeed in the first challenge, the second remains nontrivial because end users must express their criteria using the system interface. Although some criteria can be easily captured this way (such as a date interval), semantic-level criteria are more difficult to express. For these situations, formal conditions are usually combined with nonformal or fuzzy query terms (such as a query by example or a query by keywords) and IR techniques are applied.

Before describing the system design and implementation details, we first clarify the information representation space that constitutes the basis for query evaluation. Within the proposed system, incoming queries are evaluated against one or more CLE databases that, from the point-of-view of the information model, are unordered sets of image records. *Image records* refer to the combination of image data (resource) and its associated metadata. The system lets us retrieve CLE data and metadata by specifying a condition tree. So, we developed a dual database model (see Figure 2) consisting of content and metadata.

The example in Figure 3 shows two query condition trees carrying different kinds of conditions that reflect the duality of the system's information model. Both trees show conditions specifying that the metadata field date must be greater than a certain value. They also both show IR-like conditions such as a keywords or a CBIR condition using the query-by-example technique and including an example CLE image.

We designed the proposed system to evaluate queries combining keywords, CBIR, and metadata filtering. The different condition types can be combined using Boolean operators.

Querying and Metadata Interoperability

One of the system goals is providing a unified query interface for human or programmatic access, facilitating the interoperability with multiple kinds of clients and third-party applications. We designed the system's query interface for compatibility with the MPEG Query Format (MPQF, ISO/IEC 15938-12:2008) standard. However, the selection of a unified query interface is not enough to guarantee interoperability if it is not accompanied by a proper mechanism to manage metadata heterogeneity. Many different metadata formats exist for generic still-image descriptions (MPEG-7, EXIF, XMP, and so on), as well as specific standards for healthcare transactions such as Digital Imaging and Communications in Medicine (DICOM) and Health Level Seven International (HL7).

To deal with this diversity, the system's metadata model is compliant with the ISO/IEC 24800 standard (JPSearch), which is a JPEG committee initiative that aims to standardize interfaces of an abstract image-retrieval framework. Currently, our CLE retrieval system makes use of JPSearch Part 2 for representing image metadata.[8] The scope of the JPSearch Part 2 is to define a metadata model for the JPSearch framework. The model allows the use of multiple metadata formats and describes how they can be queried using the MPEG Query Format (these standards are strongly related). Consequently, our base metadata schema is the JPSearch Core Schema,[8] extended with certain fields from the DICOM standard.

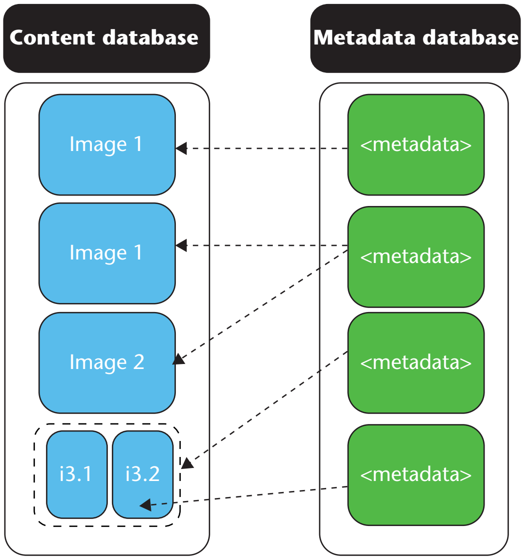

Figure 2. Dual database model. ***The system design incorporates metadata to help users express their semantic-level criteria when performing image query searches.***

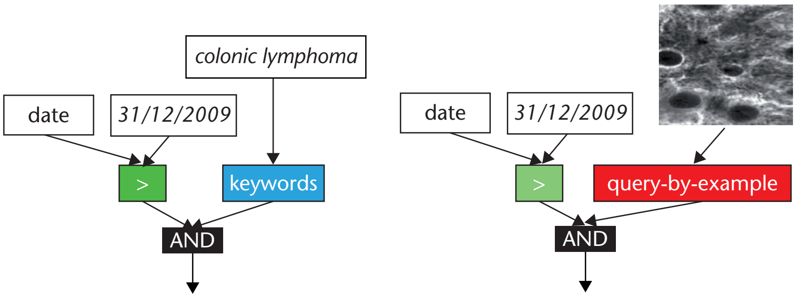

Figure 3. Example condition trees. Each tree shows information-retrieval conditions such as (a) keyword or (b) CBIR condition using the query-by-example technique.

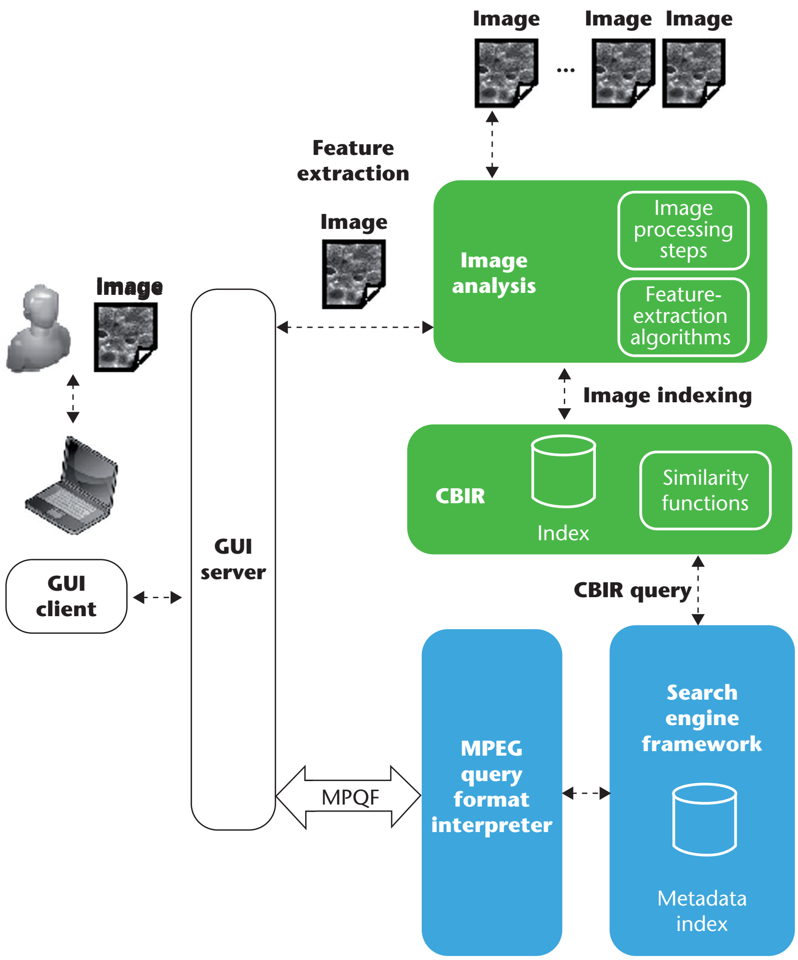

Figure 4. Overall architecture of the optical biopsy retrieval system. The system architecture consists of four main modules: image processing, CBIR index construction, the MPEG query format interpreter, and the search engine framework.

System Architecture Overview

The system architecture consists of four main modules:

1. *Image processing and analysis*. This module is applied to the off-line extraction of low-level metadata from the images in the database and to the on-the-fly extraction of the same metadata from an example image submitted by a user as a query.

2. *CBIR index construction*. This module generates an index for query-by-example search. This index implies the design of feature vectors and the selection of a similarity function.

3. *MPEG query format interpreter*. To effectively ensure interoperability with potential third-party applications, the system provides a standard query interface based on the MPEG Query Format standard.

4. *Search engine framework*. General query processor capable of solving text-based queries, CBIR queries, and combinations of both.

Figure 4 outlines the overall system architecture.

Image Processing and Analysis Module

One of the main functionalities of the proposed system is the ability to combine conventional search criteria (keywords and metadata ranges) with an example image (query-by-example paradigm) to retrieve similar, precedent cases. The latter belongs to the CBIR research area and poses several problems: the automatic extraction of low-level metadata from the images, the design of a proper feature vector, and the selection of a discriminating and efficient similarity function.

Feature-Extraction Challenge

In this work, the target images are grayscale pictures of the human colon's mucosa captured by a Pentax R EC-3870CILK. This device operates with a 475 × 475 μm field of view. Confocal frames are collected at a scan rate of 1.6 frames per second (512 × 1024 pixels) or 0.8 fps (1024 × 1024 pixels), approximating a 1,000-fold magnification on a 19-inch screen.

The colon's mucosa consists of an epithelial layer with many tubular crypts or glands going down into the tissue. The lamina propria supports the epithelium and contains blood vessels. The target images belong to en face views of the tissue, which can be captured in variable imaging-plane depth from surface to 250 μm, in contrast to the transverse sections obtained with traditional histology. Given that orientation, several colonic crypts appear visible within one image as dark circles resembling orange slices. The device lets doctors see the crypt architecture as well as the cellular and subcelluar structure (resolution 0.7 μm) because a cell is generally between 1 and 30 μm.

Our approach is only concerned with locating and measuring the crypts. In a healthy tissue, crypts appear as homogenous circles, uniformly distributed through the surface (see Figure 1). In the case of inflammation of the colon, the crypt structure changes and loses its uniformity (the intercryptal distance becomes larger), and crypt size shows greater variance. When crypt structure becomes more arbitrary, it can be a sign of *neoplasia*, or a mutation of the epithelial tissue. That can progress stepwise from adenoma to adenocarcinoma (adenoma-carcinoma sequence).

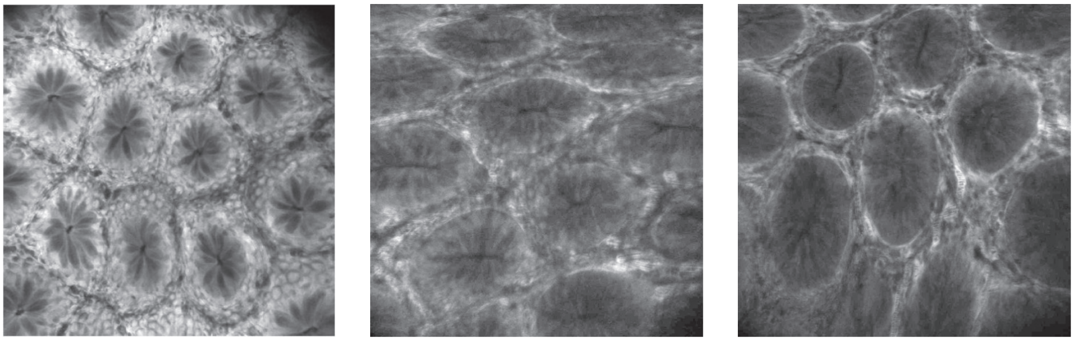

Figure 5. Three CLE images of a healthy human colon's mucosa. The degrees of fluorescence decrease over time during observations and when imaging in deeper parts of the mucosa, making it difficult to contrast various images.

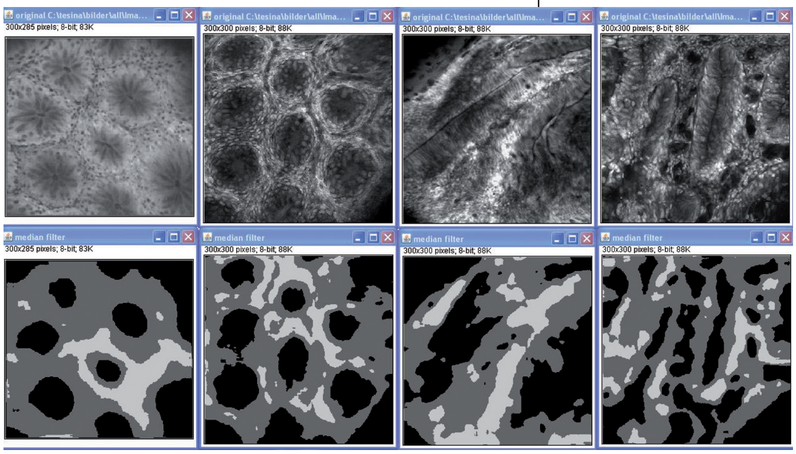

Figure 6. Image simplification stage. Four different original pictures (top) and their corresponding images after normalization and simplification (bottom).

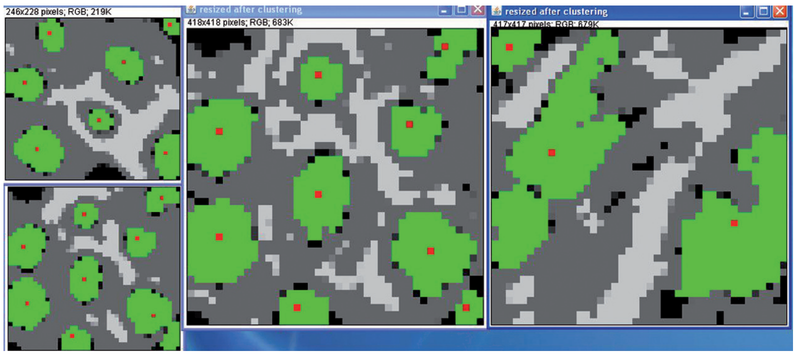

Figure 7. Feature extraction. Resulting images, three healthy (left) and one neoplastic colon (right), after the clustering process.

Another particularity of CLE images that is relevant to the feature-extraction algorithm design is the necessity to apply a contrast agent such as intravenous fluorescein or topical acriflavine. The contrast agent makes the crypts and the blood vessels (if injected) more visible, but its effects decrease during the observation and thus produce images of similarly healthy tissue with different visual aspects. This phenomenon is further augmented by a loss of fluorescence intensity when imaging in deeper parts of the mucosa (see Figure 5).

The goal of our proposed algorithm in terms of feature extraction is to quantify a set of features for similarity computation of CLE images of the colon's mucosa. The selected features should allow users to discriminate between healthy tissues, benign inflammations, and malignant lesions. In its current state, the system is unable to classify and characterize tumors.

Feature-Extraction Algorithm

Before we can measure the different features, the feature-extraction algorithm we implemented applies the local binary pattern (LBP)[6] operator to highlight the different crypts and their boundaries in the image. It is necessary to reduce the complexity of the original CLE grayscale image before we can apply this algorithm (because of its computational cost). Figure 6 shows the results of the image simplification stage.

Once an image has been properly simplified, we can calculate the LBP descriptor, the main step of the feature-extraction stage. LBP is a grayscale-invariant texture measure derived from a general definition of texture in a local neighborhood. (See related work for other LBP applications to medical image analysis.[7,8]) Once we have the LBP value for every pixel in the image, we apply a clustering algorithm to cluster pixels according to which crypt they belong to. Figure 7 shows example results. The results of this process let us extract specific features, such as the silhouette coefficient, the crypt compactness, the crypt roundness, or the intercrypt distance.

$$F(\text{image}) = \begin{pmatrix} \text{silhouette coefficient} \\ \text{silhouette coefficient variance} \\ \text{compactness} \\ \text{compactness variance} \\ \text{roundness} \\ \text{roundness variance} \\ \text{border size deviation} \\ \text{border size deviation variance} \\ \text{intercluster distance} \\ \text{intercluster distance variance} \\ \text{chi square} \end{pmatrix} = \begin{pmatrix} 0.7571 \\ 0.0048 \\ 1.0861 \\ 0.000272 \\ 0.2690 \\ 0.0142 \\ 0.0462 \\ 0.1718 \\ 4.5227 \\ 0.3272 \\ 0.1002 \end{pmatrix}$$

Figure 8. Example feature vector with common values for a healthy mucosa CLE image.

Image Indexing and Retrieval Module

Apart from the automatic extraction of low-level metadata, we still need to design a similarity function to retrieve similar images. This similarity function operates over a vector of selected features, with a composition that determines the nature of the similarity being considered. (Similarity is relative in a multidimensional space.) See Figure 8 for an example feature vector with common values for a healthy mucosa CLE image.

Before a similarity measure is computed over the feature vector, the vector should be normalized. We have applied linear scaling unit range normalization. We tested several similarity measures, such as the popular Euclidian, Manhattan, and quadratic-form distances. Our results showed that the Manhattan and Euclidian distances, in combination with the linear-scaling unit range normalization, provide better performance.

MPEG Query Format Interpreter Module

To effectively ensure the system's interoperability with potential third-party applications, we chose the MPQF standard as the search interface of the proposed architecture. MPQF is an XML-based query language that defines query formats and replies to be interchanged between clients and servers in a distributed multimedia information search-and-retrieval context. MPQF is XML-based in the sense that all MPQF instances (queries and responses) must be XML documents. Formally, MPQF is part 12 of ISO/IEC 15938, "Information Technology—Multimedia Content Description Interface" (MPEG-7). However, the query format was technically decoupled from MPEG-7, and it is now metadata neutral. Thus, MPQF is not coupled with any particular metadata standard, which lets us use it in combination with the JPSearch and DICOM metadata standards.

One of the key features of MPQF is that it is designed for articulating queries that combine the expressive styles of IR and XML data-retrieval systems (such as XQuery), embracing a range of ways to express user information needs. Traditionally, medical image-retrieval systems providing CBIR functionalities offer simple or no metadata filtering fields. This is acceptable for research prototypes because the goal of these systems is to evaluate the liability of the applied CBIR algorithms. However, the lack of powerful metadata-filtering capabilities in real-world systems severely constrains their professional use because it is necessary to specify conditions about the image metadata that appear frequently (date intervals, locations, patient's features, and so forth). This can be frustrating for users who know that the metadata are available, but the system does not provide any mechanism to combine the CBIR example image with a metadata filter.

Regarding IR-like criteria, MPQF offers a range of possibilities that include, but are not limited to, query-by-example description, query by free text, query-by-example media, query-by-example region of interest (ROI), query by feature range, query by spatial relationships, query by temporal relationships, and query by relevance feedback. Regarding XML-like criteria, MPQF offers its own XML query algebra for expressing conditions over the multimedia-related XML metadata (such as Dublin Core, MPEG-7, or any other XML-based metadata format), letting users embed XQuery expressions.

Figure 9 shows an input MPQF query asking for JPEG images taken after 15 January 2011 that have the phrase "colonic inflammation" somewhere in their metadata.

Search Engine Framework Module

Once the user requirements are formalized as an MPQF input query, the system receives it and processes it with the MPQF interpreter. This interpreter translates the query into calls to another pluggable module, the search engine framework, which is responsible for processing the different types of conditions in a query—that is, text-based queries, CBIR queries, or combinations. This design decouples a functionality from the system that can be implemented by existing IR libraries and databases. The goal is to delegate the evaluation of the query condition tree and the text-based conditions to third-party software that is also capable of interoperating with our CBIR modules.

In our implementation of the proposed architecture, we use Apache Lucene for the search engine framework. Lucene is a high-performance, full-featured text search engine library developed by the Apache Software Foundation. Because it uses its own optimized index of documents, every CLE image must be transformed into a Lucene document and indexed before any search can be conducted. Lucene is essentially a text search engine. It does not natively accept CBIR queries, but we have extended it with our CBIR modules while preserving its capability to orchestrate Boolean condition trees. This approach slightly differs from previous research,[9] which adopted the Lucene interface to process CBIR queries but completely replaces the Lucene query evaluation mechanism.

Internally, every Lucene document is shown as a collection of terms, each given a term frequency–inverse document frequency (tf-idf) weight. The Lucene scorer uses these weights to determine how similar two given documents are and thus compute their scores. Because td-idf weights are only appropriate to use with text documents and the given images are represented as a collection of features, which contain double-precision numbers, we replaced the scorer with another that uses the original values of the document fields rather than weights. Concretely, we created a class named `CBIRQuery`, which extends the Lucene `CustomScoreQuery` by reimplementing its `CustomScorer::customScore` function so that it can use our CBIR similarity function, which is based on the extracted features of every image. By using the Lucene's engine to score the documents rather than just iterating through the set of documents and manually comparing them, we have gained performance, and using Lucene's engine, we can easily perform more sophisticated queries by just combining CBIR and document-retrieval queries.

```
<MpegQuery>
  <Query>
    <Input>
      <OutputDescription>
        <ReqField>title</ReqField>
        <ReqField>date</ReqField>
      </OutputDescription>
      <QueryCondition>
        <TargetMediaType>image/jpg
        </TargetMediaType>
        <Condition xsi:type="AND">
          <Condition xsi:type="QueryByFreeText">
            <FreeText>colonic inflammation</FreeText>
          </Condition>
          <Condition xsi:type="GreaterThanEqual">
            <DateTimeField>date</DateTimeField>
            <DateValue>2011-01-15</DateValue>
        </Condition>
      </Condition>
      </QueryCondition>
    </Input>
  </Query>
```

Figure 9. Example MPEG Query Format standard input query. The query requests JPEG images that were taken after 15 January 2011 and have the phrase "colonic inflammation" somewhere in their metadata.

Results

We implemented all the modules in the proposed system in Java, with the help of certain libraries such as ImageJ (http://rsbweb.nih.gov/ij) and Lucene (http://lucene.apache.org). A GUI-enabled prototype client is available at http://dmag.ac.upc.edu/projects/biopsearch (see Figure 10). (Properly visualizing the demo requires an HTML 5 enabled browser.) The application has been also tested using smartphones in distant and isolated areas that lack pathologist support. The performance of queries involving only keywords and/or metadata directly depends on the Lucene's performance, which has been evaluated in previous work.[10]

For CBIR queries, we used a training set of 137 CLE images obtained from the Mainz University Hospital with a Pentax R EC-3870CILK device to evaluate our approach. All were JPEG images annotated using standardized metadata for JPSearch. To evaluate the retrieval quality, we divided the images into four categories (healthy, inflammatory, adenoma, and carcinoma), and we performed a k-nearest-neighbor query with the value k set to 5. We used a random image from every category as the query input.

Figure 11 shows the results of a query with a benign CLE image. On the condition that the crypts in the recorded image are clear enough

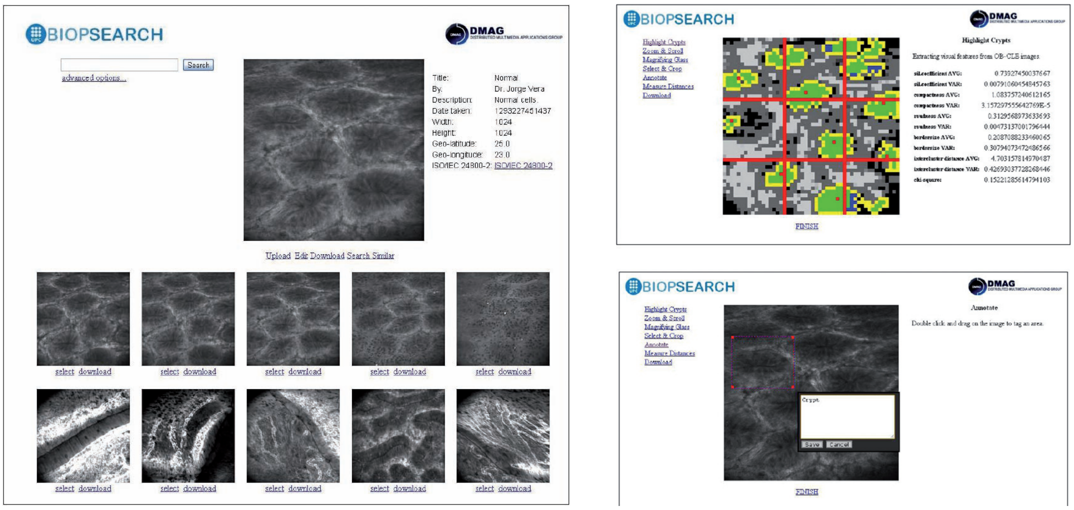

Figure 10. GUI interface of the proposed optical biopsy retrieval system.

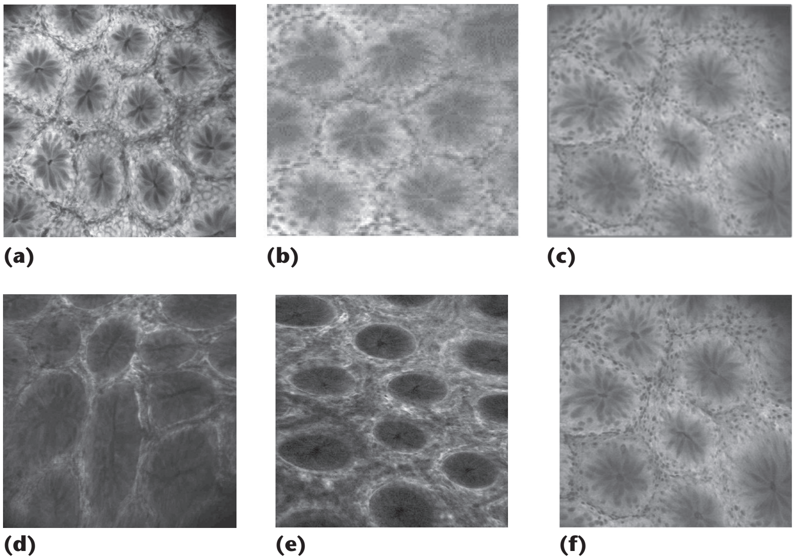

Figure 11. First five results of a CBIR query (healthy area). (a) Query image, (b) result 1, (c) result 2 (d) result 3, (e) result 4, and (f) result 5.

for the clustering algorithm to recognize, the image retrieval performance on the benign set is very good. Performing an image query with a CLE image of a carcinoma (Figure 12) also achieves good retrieval performance. Remarkably, no false positives were found—that is, no healthy images were returned for a malignant query image. The algorithm performed poorly on images where the crypts were not detected by the clustering algorithm or where the small but important irregularities were not weighted enough.

To evaluate response time, we tested the system using a Pentium D (with two Pentium 4 cores) running at 3.4 GHz. We performed the tests on the original test set and then replicated the images to create three additional repositories of increasing size:

- The *small index* contained the original test set of 137 CLE images from the Mainz University Hospital (88.1 Kbytes).
- The *medium index* contained 13,700 CLE images (5.53 Mbytes).
- The *large index* contained 1,370,000 CLE images (489 Mbytes).
- The *extra large index* contained 13,700,000 CLE images (5.34 Gbytes).

To create the image index within Lucene, we applied our feature extraction algorithm and indexed the image for each CLE image. The feature-extraction process takes between 2 and 7 seconds for a single image, and the downscaling of the input image to 50 pixels requires much of the computational time. This scaling can be optimized or even left out if the images have a standard resolution in a real operating CLE system. The indexing time is considerably faster, and the whole original test set (137 images) was indexed in less than 0.2 seconds. Table 1 shows the delay times of the indexing process (for the whole index) over the four repository sizes.

Once the index is created, solving a CBIR query involves extracting an example image's features and comparing the feature vector with each one of the feature vectors of the images in the repository. The straightforward approach would consist of iterating over all the images in the repository extracted as Lucene documents.[9] However, our tests have shown that this approach does not scale well. Table 2 shows the time necessary to traverse the whole index without actually performing the CBIR comparison.

We optimized this process by embedding the CBIR comparison within Lucene. We then tested five different queries: the metadata search only filters images given a value range for a numeric field such as weight; the similarity Euclidian search, similarity Manhattan search, similarity quadratic search, and similarity chi square search are pure CBIR queries for an example image and a given distance measure; and finally the combined search combines a metadata search with a CBIR search using the Manhattan distance.

We tested each of these queries on each repository size. Table 3 shows the numerical results of the performance evaluation. Figure 13 shows a graph of the same results. We included the document iteration times to show how our approach scales better than those that require extracting the documents from the Lucene index one by one. (The document iteration time does not include the evaluation of the distance measure.)

Future Work

Our results show that automatic discriminant feature extraction from CLE images is feasible in terms of precision, recall, and response time. It could be even faster if the CLE images are obtained with a standard resolution. However, we must still address the system's liability from the perspective of the rigorous standards in the healthcare domain. The system must be evaluated against a more complete test suite, after which we can address reimbursement and liability issues. Nevertheless, from a clinical point of view, the proposed image retrieval system may greatly facilitate CLE database management and assistance in diagnosis.

In addition, feedback from medical experts has revealed that the results of the feature-extraction step, taken as a suggested automatic diagnosis, are interesting by themselves. So, we

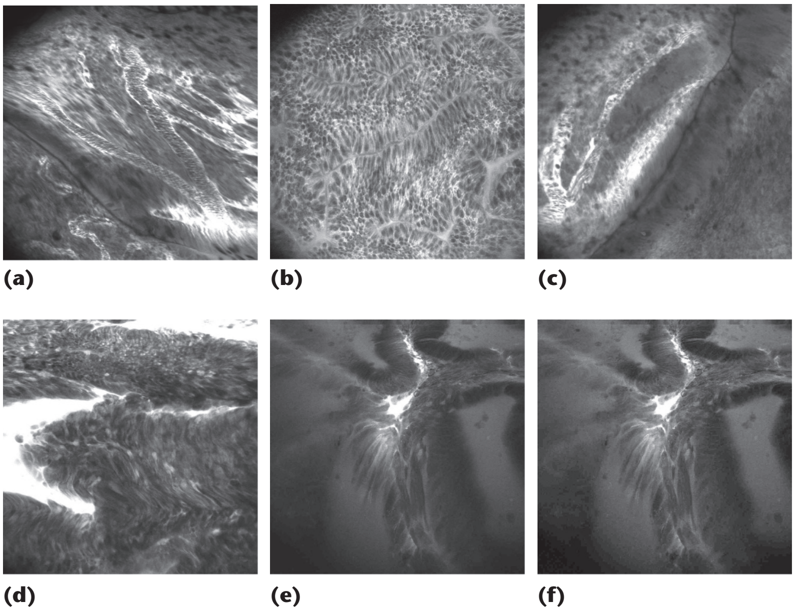

Figure 12. First five results of a CBIR query (carcinoma). (a) Query image (b), result 1 (c), result 2 (d), result 3, (e) result 4, and (f) result 5.

Table 1. Performance results during index creation.

Index	Indexing time (ms)
Small	187
Medium	2,625
Large	275,843
Extra large	2,928,406

Table 2. Document traversal times in Lucene.

Index	Document traversal time (ms)
Small	63
Medium	328
Large	18,500
Extra large	654,515

Table 3. System performance results.

	Index*			
Query type	Small	Medium	Large	Extra large
Metadata search	15	15	171	1,516
Similarity Euclidian search	16	62	4,234	40,671
Similarity Manhattan search	15	47	3,062	27,969
Similarity quadratic search	31	937	85,703	–
Similarity chi square search	16	78	4,985	47,719
Combined search	16	62	2,718	27,969

* All values are in milliseconds.

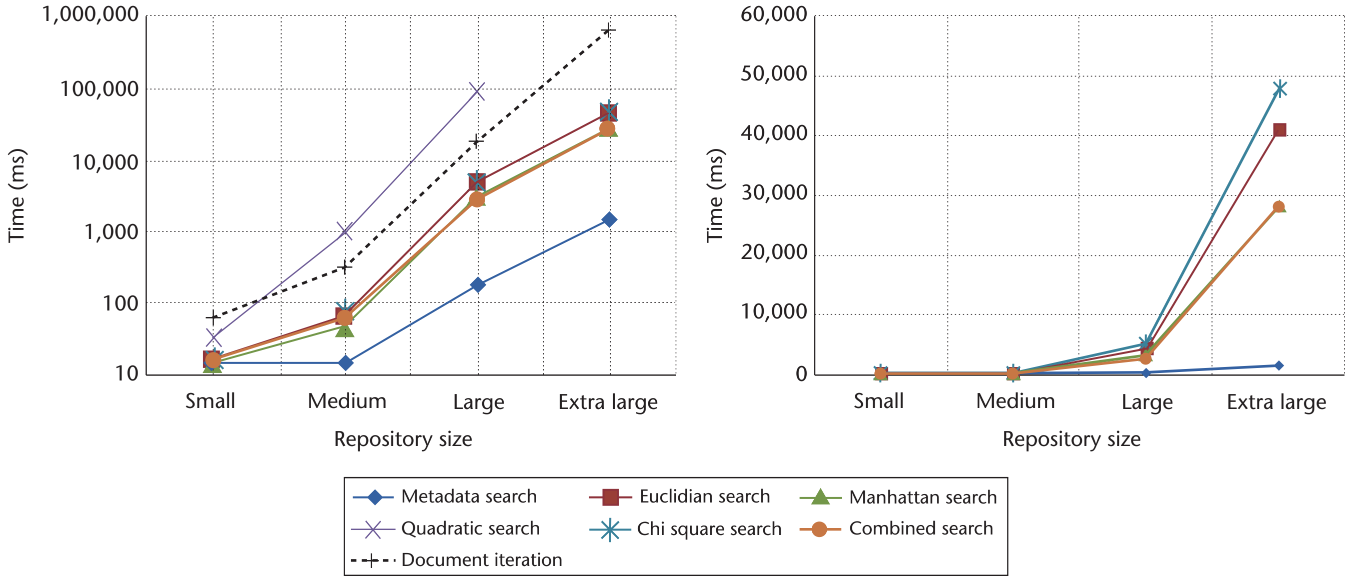

Figure 13. Experimental response times. (a) Different search times in logarithmic scale. (b) Linear scale without quadratic search.

are currently applying machine learning techniques to automatically classify images into normal, inflammatory, hyperplastic, and neoplastic tissue. Suggesting a diagnosis presents multiple challenges because we must guarantee that no false positives appear for the "normal" category—that is, abnormal tissue misclassified as normal. There are also challenges in terms of performance because the diagnosis must be provided in real time. Future goals include further retrieval evaluation in coordination with medical experts from Mainz University Hospital and integration of a test prototype within their CLE equipment. **MM**

Acknowledgments

This work has been partly supported by the Spanish government (under grant TEC2011-22989) and by the UNESCO chair of telemedicine and the telemedicine chair of Telefonica SA.

References

1. M. Döller et al., "The MPEG Query Format: On the Way to Unify the Access to Multimedia Retrieval Systems," *IEEE MultiMedia,* vol. 15, no. 4, 2008, pp. 82–95.
2. *ISO/IEC 24800-2:2011 Information Technology – JPSearch – Part 2: Registration, Identification, and Management of Schema and Ontology,* Int'l Organization for Standardization, 2011.
3. R. Kiesslich, P.R. Galle, and M.F. Neurath, *Atlas of Endomicroscopy,* Springer, 2008.
4. O. Ferrer-Roca, "Telepathology and Optical Biopsy," *Int'l J. Telemedicine and Applications,* vol. 2009, article ID 740712, doi:10.1155/2009/740712.
5. K.B. Dunbar et al., "Confocal Laser Endomicroscopy in Barrett's Esophagus and Endoscopically Inapparent Barrett's Neoplasia: A Prospective, Randomized, Double-Blind, Controlled, Crossover Trial," *Gastrointestinal Endoscopy,* vol. 70, no. 4, 2009, pp. 645–654.
6. M. Pietikainen, T. Ojala, and T. Maenpaa, "Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *Proc. 6th European Conf. Computer Vision* (ECCV 00), Springer-Verlag, 2000, pp. 404–420.
7. L. Nanni, A. Lumini, and S. Brahnam, "Local Binary Patterns Variants as Texture Descriptors for Medical Image Analysis," *Artificial Intelligence in Medicine,* vol. 49, no. 2, 2010, pp. 117–125.
8. X. Xu and Q. Zhang, "Medical Image Retrieval Using Local Binary Patterns with Image Euclidean Distance," *Proc. Int'l Conf. Information Eng. and Computer Science,* IEEE Press, 2009, pp. 1–4.
9. M. Lux and S.A. Chatzichristofis, "Lire: Lucene Image Retrieval: An Extensible Java CBIR Library," *Proc. 16th ACM Int'l Conf. Multimedia* (MM 08), ACM Press, 2008, pp. 1085–1088.
10. Y. Jing, C. Zhang, and X. Wang, "An Empirical Study on Performance Comparison of Lucene and Relational Database," *Proc. Int'l Conf. Communication Software and Networks,* IEEE CS Press, 2009, pp. 336–340.

Ruben Tous an associate professor in the Department of Computer Architecture at Universitat Politècnica de Catalunya, BarcelonaTech (UPC). His scientific work focuses on algorithms and data structures, knowledge representation and reasoning for multimedia understanding, multimedia databases and

query languages, and multimedia information retrieval. Tous has a PhD in computer science and digital communications from Universitat Pompeu Fabra, Spain. Contact him at rtous@ac.upc.edu.

Jaime Delgado is a professor in the Computer Architecture Department and the founder and head of the Distributed Multimedia Applications Group (DMAG) at Universitat Politècnica de Catalunya, BarcelonaTech (UPC). His research interests include multimedia applications, privacy, metadata interoperability, multimedia search, and digital rights management. Delgado has a PhD in telecommunications engineering from UPC. Contact him at jaime.delgado@ac.upc.edu.

Thomas Zinkl is a doctoral student in the Department of Computer Architecture at Universitat Politècnica de Catalunya, BarcelonaTech (UPC). His research interests include multimedia information retrieval and image analysis. Zinkl has an MS in computer science from Universität Passau. Contact him at zinklthomas@gmail.com.

Pere Toran is a researcher in the Distributed Multimedia Applications Group (DMAG) at Universitat Politècnica de Catalunya, BarcelonaTech (UPC). His scientific work focuses on information retrieval in medical applications, query languages, and digital rights management. Toran has an MS in information technology from UPC. Contact him at ptoran@ac.upc.edu.

Gabriela Alcalde a doctoral student in the Department of Computer Architecture at Universitat Politècnica de Catalunya, BarcelonaTech (UPC). Her research interests include computer networks, security, privacy, access control, and digital rights management. Alcalde has an MS in computer science in from Instituto Tecnológico de Orizaba. Contact her at gabriela@ac.upc.edu.

Martin Goetz is an associate professor, head of the clinical inflammatory bowel disease unit, and second head of interdisciplinary endoscopy at Mainz University Hospital, Germany. His research and clinical work focuses on advanced diagnostic and therapeutic endoscopy and inflammatory bowel diseases. Goetz received specialist training as an internist and gastroenterologist at the Ludwig-Maximilians-University Munich and the Johannes Gutenberg-University Mainz. Contact him at mgoetz@uni-mainz.de.

Olga Ferrer-Roca is the UNESCO Chair of Telemedicine and chair of pathology at the University of La Laguna, Tenerife, Spain. Her research interests include innovation in medicine, telemedicine, and oncology. Ferrer-Roca has a PhD in medicine from the Central University of Barcelona. Contact her at catai@teide.net.

Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.